**PREWORK**

**SESIÓN 08**

**Objetivo:**

En esta sesión veremos las bases de procesamiento de imágenes para Machine Learning.

**Pre-requisitos (Técnicos):**

En esta sesión necesitaremos que instales OpenCV. Esta librería (Open Computer Vision) es una herramienta que te permitirá manipular imágenes digitales por medio de algoritmos y programas. Muchas herramientas de edición de imágenes como GIMP o Photoshop trabajan con OpenCV, y puedes instalarlo con el siguiente comando:

*pip3 install opencv-python*

**Pre-requisito 1: Datos no estructurados y no estructurados.**

Para comenzar con este pre-requisito, primero debemos de entender una crucial diferencia entre los datos que hemos estado manejando: Casi todos los sistemas de Machine Learning que vimos en este módulo trabajan con ***datos estructurados***, los cuales son datos que han tenido cierto procesamiento previo, cierto orden y quizás vienen de una base de datos o un archivo JSON.

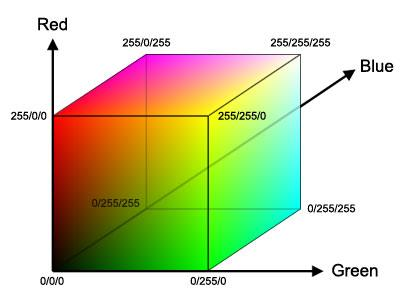
Estos datos con estructura, orden y preprocesamiento son muy buenos para ser tratados por Machine Learning, sin embargo, se estima que el 80% de los datos existentes en el mundo no tienen tanta suerte: Muchas veces los datos pueden provenir de fuentes análogas como cuadernos escritos a mano alzada, pueden estar capturados en imágenes digitales sin ningún contexto, o bien, pueden venir de sistemas que se programaron hace años, y no hay una manera sencilla de entender como es que se capturaron o que sentido tienen.

Estos datos se le conocen como ***datos no estructurados***, y son un obstáculo para que puedan aprovecharse de la mejor manera. Podemos ver una comparativa entre datos estructurados y datos no estructurados en esta tabla:

|  |  |  |
| --- | --- | --- |
|  | Estructurados | No estructurados |
| Fuente | Bases de datos, archivos | Imágenes, sensores, audios |
| Dificultad de uso | Fácil | Compleja |
| Uso en machine learning | Alto | Alto |
| Probabilidad de aparición | 20% | 80% |

**Pre-requisito 2: Píxeles y teoría del color.**

Cuando las imágenes digitales comenzaron, el concepto de píxel nació con ellas. Un píxel es una representación a color de tres valores: Un color Rojo, un Color Verde y un color Azul. En general, la representación de color va de 0 a 255 para cada uno de estos tres canales, y su combinación puede generar lo que le llamamos color verdadero: mas de 15 millones de colores diferentes que podemos observar.



Sin embargo, si tu utilizas pintura roja, verde y azul sobre un lienzo, ¿porque no puedes obtener un color blanco uniforme? La respuesta se encuentra en la forma en la que el color trabaja: Cuando el color va acompañado de luz, como es la pantalla de tu computadora, los colores se suman entre si hasta obtener blanco. Esto se le conoce como *color por adición*.

Sin embargo, cuando pintas en un lienzo en blanco, cada pintura resta “color” al lienzo: La pintura azul le quita todos los demás colores al lienzo blanco, excepto el azul, por ello se ve de ese color. Es por ello que la pintura actúa con un fenómeno llamado *color por substracción.* Es por ello que nosotros trabajaremos con color por adición, debido a que las imágenes digitales suelen necesitar el espacio de color RGB y se despliegan en un monitor para su uso.

**Pre-requisito 3: Visión computacional y brecha semántica.**

La visión computacional nace con la necesidad de hacer que las computadoras puedan entender el mundo como lo entendemos los humanos por medio de imágenes o video (que no es otra cosa que cuadros de imágenes, presentados uno tras otro). Si bien es cierto que la visión la poseen también los animales, la visión humana es una de nuestras principales fuentes de información, y nuestro objetivo es darles a las máquinas la misma capacidad de entendimiento del ambiente por medio de imágenes.

Compañías como Tesla Motors invierten muchísimo dinero y esfuerzo para que los automóviles puedan entender como funciona el mundo a través de un volante, una cámara y unos sensores. Sin embargo, como puedes imaginar, una imagen es solamente un conjunto de píxeles. ¿Como es que un montón de pixeles y valores numéricos podría tener algo de sentido para una computadora?

Esta brecha entre un puñado de números y el entendimiento puro del ambiente que le rodea a una máquina se le conoce como *brecha semántica*. El objetivo de los investigadores en visión computacional es reducir la brecha semántica de tal manera que la capacidad de entendimiento de una computadora por medio de fotogramas sea exactamente igual de buena que la de un ser humano.

Durante esta sesión, utilizaremos varias herramientas que, aunadas con machine learning, nos ayudarán a cerrar poco a poco esta brecha semántica para crear sistemas mas inteligentes, que puedan realizar labores que solamente habían estado reservados para los seres humanos.

**QUIZ**

**¿Qué significa “color por substracción”?**

1. Cuando un color se mezcla con otro color
2. Cuando un color es sobreescrito por otro color
3. Cuando un color quita la opacidad de otro color
4. **Cuando un color quita la luz de otro color**

**¿Qué es la brecha semántica?**

1. La capacidad de las computadoras de entender su ambiente.
2. El espacio que hay entre la inteligencia natural y la inteligencia artificial.
3. **El espacio que existe entre el entendimiento humano y el artificial.**
4. El tiempo de procesamiento de un algoritmo y un ser humano.

**¿Por qué nos interesa trabajar con datos no estructurados?**

1. Porque los datos no estructurados no requieren procesamiento previo.
2. Porque los datos no estructurados son más fáciles.
3. **Porque la cantidad de problemas con datos no estructurados es mucho mayor.**
4. Porque los datos estructurados son muy abundantes en todos lados.

**¿Por qué usamos el espacio de color RGB?**

1. **Porque es un estándar de color por adición.**
2. Porque es un estándar de color por substracción.
3. Porque es una manera de innovar en los colores.
4. Porque el espacio de color CYMK está ocupado.